Journal of the Korean Association of Geographic Information Studies, 24(1) 2021, pp.1-11 ISSN 1226-9719(Print)
https://doi.org/10.11108/kagis.2021.24.1.001 ISSN 2287-6952(Online)

Landsat-8 위성영상 기반 수분지수 및 기계학습을 활용한 대구광역시의 지표수 탐지*

정윤재[1※] · 김경섭[2] · 박인선[2] · 정연인[3]

Detection of Surface Water Bodies in Daegu Using Various Water Indices and Machine Learning Technique Based on the Landsat-8 Satellite Image*

Yun-Jae CHOUNG[1※] · Kyoung-Seop KIM[2]
In-Sun PARK[2] · Youn-In CHUNG[3]

요 약

위성영상을 활용한 하천, 습지, 호수 등 지표수 객체의 탐지는 해당 지역의 수자원 관리 및 조사 업무에 효율적으로 활용될 수 있다. 본 연구에서는 원격탐사 분야에서 물을 탐지하기 위해 제공하는 수분지수(Water Index)와 영상으로부터 객체를 인식하는 데 폭넓게 활용되는 기계학습(Machine learning) 기법을 대구광역시를 촬영한 Landsat-8 위성영상에 개별적으로 적용하여 하천, 호수 등 다양한 지표수 객체를 탐지하고 그 결과를 비교하였다. 우선 Landsat-8 위성영상의 다중분광 밴드로부터 NDWI(Normalized Difference Water Index), MNDWI(Modified Normalized Difference Water Index) 영상을 생성하였고, 임계치를 적용하여 개별 영상으로부터 물과 그 외 지역을 구분할 수 있는 이진 영상(Binary image)을 제작하였다. 그리고 기계학습 기법인 SVM(Support Vector Machine)을 동일 위성영상에 적용하여 토지 피복 영상을 제작하고 이로부터 이진 영상을 제작하였다. 최종적으로 100개의 검사점(Checkpoints)을 사용하여 세 이진 영상으로부터 지표수 탐지를 위한 정확도를 오차 행렬을 활용하여 계산하였다. 그 결과, MNDWI 영상으로부터 제작된 이진 영상의 정확도(84%)가 NDWI 영상으로부터 제작된 이진 영상의 정확도(94%)와 SVM에 의해 제작된 이진 영상의 정확도(96%)에 비해 낮았으며, 모든 이진 영상에서 그림자 등의 원인으로 인해 일부 육지 분류 결과가 지표수 객체로 오분류되었다.

주요어 : Landsat-8 위성영상, 지표수, NDWI, MNDWI, 기계학습, SVM

2020년 12월 10일 접수 Received on December 10, 2020 / 2020년 12월 18일 수정 Revised on December 18, 2020 / 2020년 12월 24일 심사완료 Accepted on December 24, 2020

* 본 연구는 국토교통부/국토교통과학기술진흥원의 지원으로 수행되었음 (과제번호 20DPIW-C153746-02)

1 ㈜지오씨엔아이 공간정보기술연구소 연구소장 Director, Geospatial Research Center, GEO C&I., Ltd.
2 ㈜지오씨엔아이 공간정보기술연구소 연구원 Researcher, Geospatial Research Center, GEO C&I., Ltd.
3 계명대학교 토목공학과 교수 Professor, Dept. of Civil Engineering, Keimyung University
※ Corresponding Author E-mail : choung12osu@gmail.com

ABSTRACT

Detection of surface water features including river, wetland, reservoir from the satellite imagery can be utilized for sustainable management and survey of water resources. This research compared the water indices derived from the multispectral bands and the machine learning technique for detecting the surface water features from he Landsat-8 satellite image acquired in Daegu through the following steps. First, the NDWI(Normalized Difference Water Index) image and the MNDWI(Modified Normalized Difference Water Index) image were separately generated using the multispectral bands of the given Landsat-8 satellite image, and the two binary images were generated from these NDWI and MNDWI images, respectively. Then SVM(Support Vector Machine), the widely used machine learning techniques, were employed to generate the land cover image and the binary image was also generated from the generated land cover image. Finally the error matrices were used for measuring the accuracy of the three binary images for detecting the surface water features. The statistical results showed that the binary image generated from the MNDWI image(84%) had the relatively low accuracy than the binary image generated from the NDWI image(94%) and generated by SVM(96%). And some misclassification errors occurred in all three binary images where the land features were misclassified as the surface water features because of the shadow effects.

KEYWORDS : Landsat-8 Satellite Image, Surface Water, Normalized Difference Water Index, Modified Normalized Difference Water Index, Machine Learning, Support Vector Machine

서 론

지표수(Surface water)는 지구 표면에 있는 물을 의미하며, 일반적으로 바다를 제외한 하천, 호수, 저수지 등의 물을 통틀어 가리킨다(ScienceAll, 2020). 위성영상 등 원격탐사 자료를 활용한 지표수 조사는 개별 지표수의 현황 및 특성을 현장에 방문하지 않고 파악할 수 있다는 장점이 있어 최근 국내외에서 폭넓게 활용되고 있다(Jo, 2012; Kim *et al.*, 2017; Kim and Lee, 2018; Hwang *et al.*, 2018).

광학 센서를 활용하여 취득한 위성영상의 다중분광 밴드는 밴드별 특성 및 파장에 따라 특정 객체에 민감하게 반응한다 (Jensen, 2006). 국내외 원격탐사 분야 연구자들은 이러한 위성영상의 밴드별 특징을 활용하여 하천, 호수, 습지 등 수자원 객체를 조사 및 탐지하기 위한 수분지수를 개발하였다. McFeeters(1996)는 Landsat-5 TM(Thematic Mapper) 위성영상의 가시광선 Green 밴드와 근적외선 밴드를 활용하여 정규수분지수(NDWI, Normalized Difference Water Index)를 제안하였다. Hu(2006)은 Landsat-5 TM 위성영상의 근적외선 밴드와 중적외선 밴드를 활용하여 수정된 정규수분지수(MNDWI, Modified Normalized Difference Water Index)를 제안하였다. 수분비율지수는 Landsat-7 ETM+(Enhanced Thematic Mapper) 위성영상의 가시광선 Green, Red 밴드와 근적외선 및 중적외선 밴드를 활용하여 만들어진 수분지수로서, Shen and Li(2010)에 의해 제안되었다.

위성영상의 밴드별 특성에 기반해 제작된 다양한 수분지수를 활용한 수자원 조사 연구는 국

내외 연구자들에 의해 진행되었다. Zhai *et al.*(2015)는 Landsat 위성영상 기반 수분지수를 활용하여 중국 Tianjin 지역의 수자원 조사 연구를 수행하였다. Acharya *et al.*(2018)은 Landsat 위성영상 기반 수분지수를 활용하여 네팔의 지표수 조사 연구를 수행하였다.

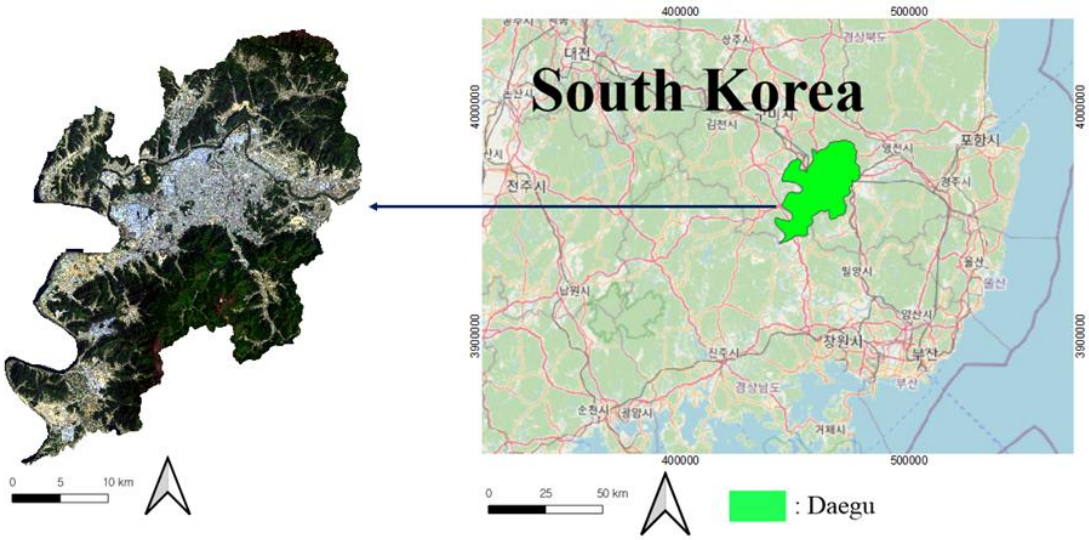
기계학습(Machine learning)은 인공지능의 한 분야로서 수학적 모델을 사용하여 데이터 내의 패턴을 식별하고, 이러한 패턴을 사용하여 예측할 수 있는 데이터 모델을 만드는 알고리즘을 의미한다(Microsoft, 2020). 원격탐사 분야에서는 위성영상 분석에 기계학습을 적용한 지표수 탐지 연구가 진행되었다. Jo(2012)는 RapidEye 위성영상에 기계학습 기법을 적용하여 위성영상에서 하천(낙동강)을 탐지하였다. Acharya *et al.*(2019)는 Landsat 위성영상에 기계학습 알고리즘을 적용하여 네팔의 지표수 조사 연구를 수행하였다.

위와 같이 위성영상에서 지표수를 탐지하기 위해 수분지수 또는 기계학습 기법을 개별적으로 적용한 연구는 진행되었으나, 두 접근법의 비교에 관한 연구는 상대적으로 부족하였다. 본 연구에서는 테스트베드(Test bed)를 촬영한 위성영상에 수분지수와 기계학습을 개별적으로 적용하여 주어진 위성영상으로부터 하천, 호수 등 다양한 지표수를 탐지하고 그 정확도를 산출함으로써 두 접근법을 비교하였다.

연구 방법

1. 연구 대상 지역 및 범위

본 연구에서는 대상 지역 구름 등 기상에 의한 영상 왜곡 현상을 최소화하기 위하여 영상 내의 구름의 양이 10% 이하인 상태에서 2017년 4월 29일에 대구광역시를 촬영한 Landsat-8 위성영상을 USGS(United States Geological Survey, 2020)로부터 획득하여 사용하였다. 그림 1에서는 2017년 4월 29일에 대구광역시를 촬영한 Landsat-8 위성영상(그림 1(a))과 OSM(Open Street Map, 2020)에서 확인할 수 있는 대구광역시의 위치(그림 1(b))를 확인할 수 있다.



(a) (b)

FIGURE 1. Landsat-8 satellite image acquired in Daegu, South Korea (a) shown in Open Street Map (b)

TABLE 1. Metadata of Landsat-8 satellite image used in this research(USGS, 2020)

	Landsat-8 satellite image used in this research
Spatial resolution	30m
Horizontal datum	WGS(World Geodetic System) 84 / UTM(Universal Transverse Mercator) zone 52N
Bands	Band 1:0.43-0.45㎛ Band 2:0.45-0.5㎛ Band 3:0.53-0.59㎛ Band 4:0.64-0.67㎛ Band 5:0.85-0.88㎛ Band 6:1.57-1.65㎛ Band 7:2.11-2.29㎛ Band 8:0.50-0.68㎛ Band 9:1.36-1.38㎛ Band 10:10.6-11.19㎛ Band 11:11.5-12.51㎛

본 연구에서 사용한 Landsat-8 위성영상의 메타데이터는 표 1에서 확인할 수 있다.

2. 연구 방법

본 연구에서는 원격탐사에서 주로 사용되는 수분지수인 NDWI, MNDWI와 기계학습 기법인 SVM(Support Vector Machine)을 주어진 Landsat-8 위성영상에 개별적으로 적용하여, 영상으로부터 물과 육지(Land)를 구분하는 이진 영상(Binary image)을 제작하고 지표수 객체를 탐지하였다. 그림 2는 본 연구에서 수분지수와 기계학습 방법으로 Landsat-8 위성영상을 활용하여 대구광역시의 다양한 지표수를 탐지하는 과정을 순서도로 나타낸 것이다.

본 연구에서 사용한 수분지수는 NDWI(Normalized Difference Water Index)와 MNDWI(Modified Normalized Difference Water Index)이며, 각 수분지수를 제작하는 원리는 식 1과 2에서 확인할 수 있다(McFeeters, 1996; Hu, 2006).

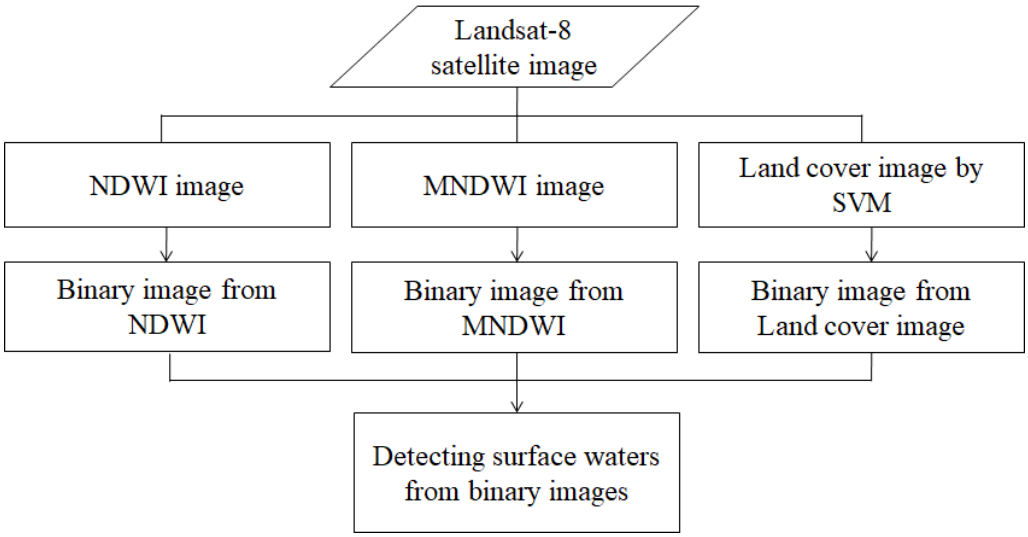


FIGURE 2. Flowchart showing the procedure of detecting surface waters from the given Landsat-8 satellite image

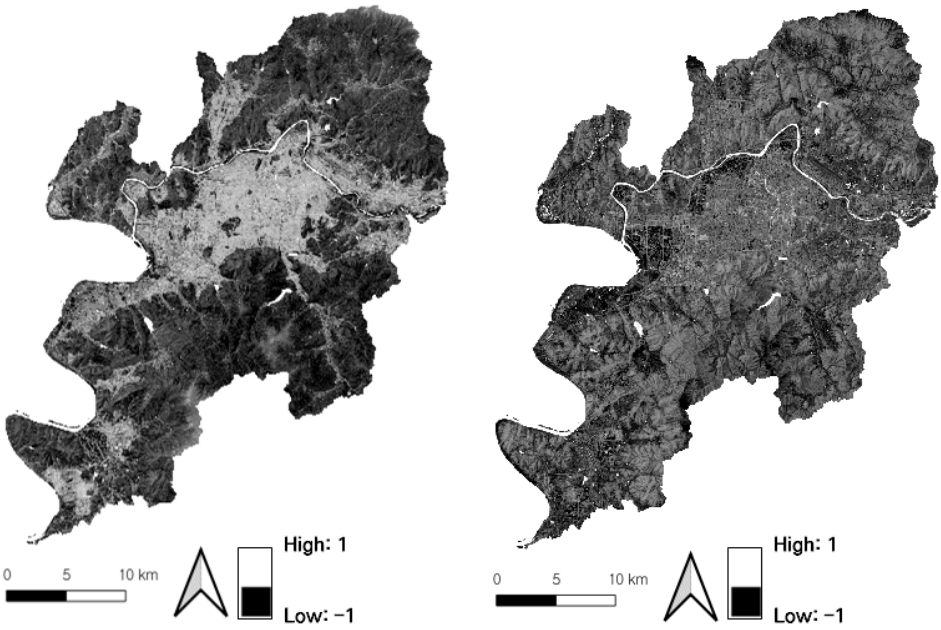
$$NDWI = \frac{Band3 - Band5}{Band3 + Band5} \quad (1)$$

$$MNDWI = \frac{Band3 - Band6}{Band3 + Band6} \quad (2)$$

식 1, 2에서 Band 3은 Green 밴드, Band 5는 근적외선(NIR, Near Infrared) 밴드, Band 6은 Landsat-8 위성영상의 단파장 적외선(SWIR, Short-Wave Infrared) 밴드 1을 의미한다. 그림 3은 주어진 Landsat-8 위성영상의 다중분광 밴드를 활용하여 제작한 대구광역시의 NDWI 및 MNDWI 영상을 보여준다.

그림 3에서 보는 바와 같이, 주어진 Landsat-8 위성영상의 다중분광 밴드를 활용하여 제작한 NDWI 및 MNDWI 영상에서 밴드별 연산에 의해 픽셀의 최댓값 및 최솟값은 각각 1 및 -1의 값을 가진다. 일반적으로 NDWI 및 MNDWI 영상에서 픽셀의 밝기 값이 1에 가까울수록 물 지역(수체)을 의미하고 -1에 가까울수록 식생, 나지, 도심 등 육지 지역을 의미하며, 픽셀의 밝기 값이 0 이상이면 수체로 정의한다(McFeeters, 1996; Hu, 2006; Choung and Jo, 2016). 본 연구에서는 그림 3의 NDWI 및 MNDWI 영상에서 밝기 값이 0 이상인 픽셀은 수체, 그 외에는 식생, 나지, 도심 지역을 포함한 육지로 정의하였고, NDWI 및 MNDWI 영상으로부터 각각 이진 영상(Binary image)을 제작하였다(그림 4).

SVM은 패턴 인식, 객체 탐지 등 폭넓은 분야에서 활용되는 기계학습 알고리즘으로서, 다차원의 집단에서 초평편(Hyperplane)을 적용하여 n개의 클러스터(Cluster)를 구분한다(Choung and Jo, 2017; Choung and Kim, 2019). 본 연구에서는 주어진 Landsat-8 위성영상의 공간 해상도(30m)를 고려하여 위성영상으로부터 대구광역시의 주요 토지 피복인 물, 도심, 나지 및 식생 지역을 구분하는 토지 피복 영상을 제



(a) (b)

FIGURE 3. NDWI image(a) and MNDWI image(b) derived from the Landsat-8 satellite image acquired in Daegu, South Korea

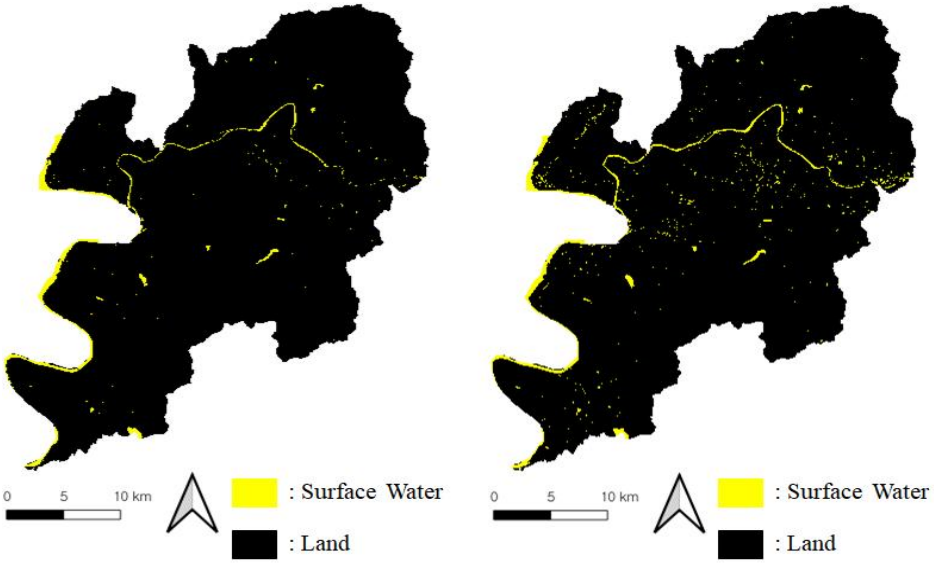
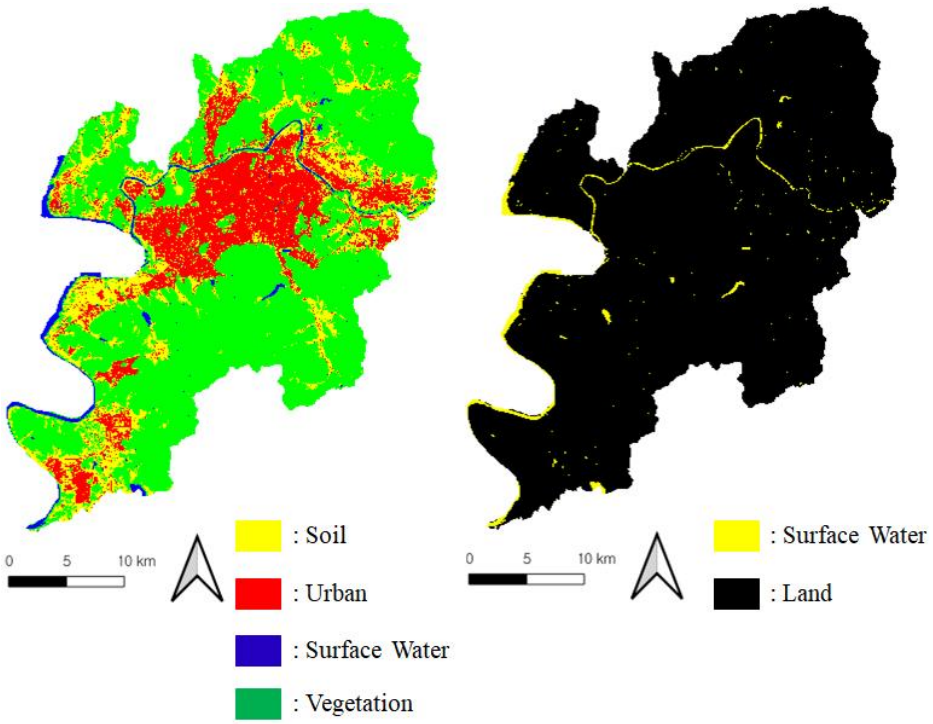


FIGURE 4. Binary images from the NDWI image(a) and MNDWI image(b)



(a) (b)

FIGURE 5. Land cover image generated by SVM(a) and the binary image generated from the land cover image(b)

작하였다. Choung and Kim(2019)이 수행했던 연구와 동일하게 Landsat-8 위성영상으로부터 각 토지 피복(물, 도심, 나지 및 식생)별로 1,000개의 훈련 자료(Training sample)를 취득한 후 SVM 알고리즘을 적용하여 토지 피복 영상을 제작하였고(그림 5(a)), 토지 피복 영상으로부터 물을 제외한 나머지 지역을 합쳐 물과 육지가 구분이 가능한 이진 영상을 제작하였다(그림 5(b)).

결과 및 고찰

본 연구에서는 주어진 Landsat-8 위성영상에 NDWI, MNDWI 및 SVM을 적용하여 제작한 개별 이진 영상(그림 4(a), 4(b) 및 5(b))의 정확도를 검증하기 위해 총 100개의 검사점(지표수 지역에 30개, 육지 지역에 70개)을 이용하여 개별 이진 영상에 대한 오차행렬(Error matrix)을 작성하였다. 오차행렬은 검사점 위치에서 이진 영상의 픽셀 분류 결과와 Landsat-8 위성영상의 토지 피복 현황을 활용하여 검증한 결과를 기반으로 정확도를 수치화한 도표로서 원격탐사/인공지능 기법을 활용하여 제작한 영상 기반 토지 피복 영상의 정확도 검증에 주로 활용된다(Moon *et al.*, 2020). 그림 6은 개별 이진 영상의 정확도를 검증하기 위해 이용한 100개의 검사점의 위치를 보여주며, 오차 행렬을 사용하여 산출한 개별 이진 영상의 정확도는 표 2~4에서 확인할 수 있다.

표 2~4에서 보는 바와 같이 MNDWI 영상으로부터 생성된 이진 영상의 정확도(84%) 및 Kappa 계수(0.65)는 NDWI 영상으로부터 생성

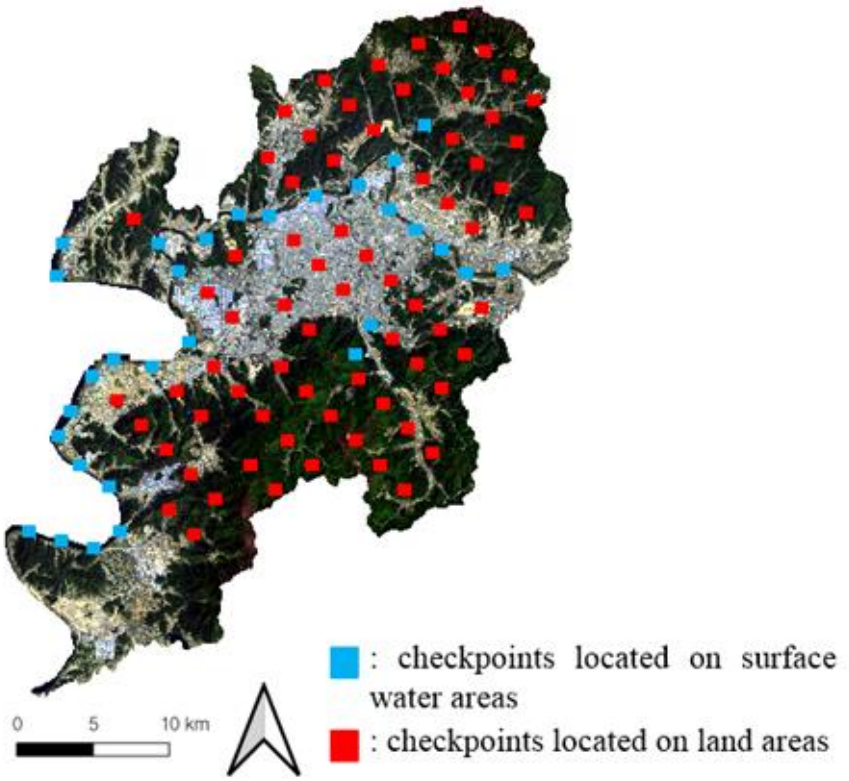


FIGURE 6. Checkpoints located on land and surface water areas shown in the given Landsat-8 satellite image

TABLE 2. Accuracy of the binary image generated from NDWI

Class	Surface water (Reference)	Land (Reference)	Total	User' s accuracy
Surface water (Classification)	27	3	30	90%
Land (Classification)	3	67	70	95.7%
Total	30	70	100	
Producer' s accuracy	90%	95.7%		94%(Overall)
Kappa coefficient		0.86		

TABLE 3. Accuracy of the binary image generated from MNDWI

Class	Surface water (Reference)	Land (Reference)	Total	User' s accuracy
Surface water (Classification)	27	13	40	67.5%
Land (Classification)	3	57	60	95%
Total	30	70	100	
Producer' s accuracy	90%	81.4%		84%(Overall)
Kappa coefficient		0.65		

TABLE 4. Accuracy of the binary image generated by SVM

Class	Surface water (Reference)	Land (Reference)	Total	User' s accuracy
Surface water (Classification)	28	2	30	93.3%
Land (Classification)	2	68	70	97.1%
Total	30	70	100	
Producer' s accuracy	93.3%	97.1%		96%(Overall)
Kappa coefficient		0.90		

된 이진 영상의 정확도(94%) 및 Kappa 계수(0.86)와 SVM에 의해 생성된 이진 영상의 정확도(96%) 및 Kappa 계수(0.90)에 비해 낮았고, 이는 MNDWI 기반 이진 영상이 NDWI 및 SVM 기반 이진 영상에 비해 육지(Land)와 지표수(Surface water)를 구분하는데 상대적으로 많은 오류가 발생한다는 사실을 의미한다. 그림 7에서 보는 바와 같이 MNDWI 영상으로부터 생성된 이진 영상(그림 7(c))에서 다른 두 이진 영상(그림 7(b) 및 7(d))에 비해 상대적으로 더 많은 육지 분류 결과가 지표수 객체로 오분류(Misclassification)되었다. 물론 그림 7(b) 및 7(d)와 같이 나머지 두 이진 영상에서도 일부 육지 분류 결과가 지표수 객체로 오분류되어 있었으며, 육안으로 확인하였을 때 그림자 등이 원인으로 파악되었다.

결 론

본 연구에서는 대구광역시를 촬영한 Landsat-8 위성영상에 NDWI 및 MNDWI와 같은 수분지수 및 SVM과 같은 인공지능 기법을 개별적으로 적용하여 하천, 호수 등 다양한 지표수를 탐지하였고 정확도를 비교하였다. 본 연구를 통해 실험한 결과 수분지수인 NDWI와 인공지

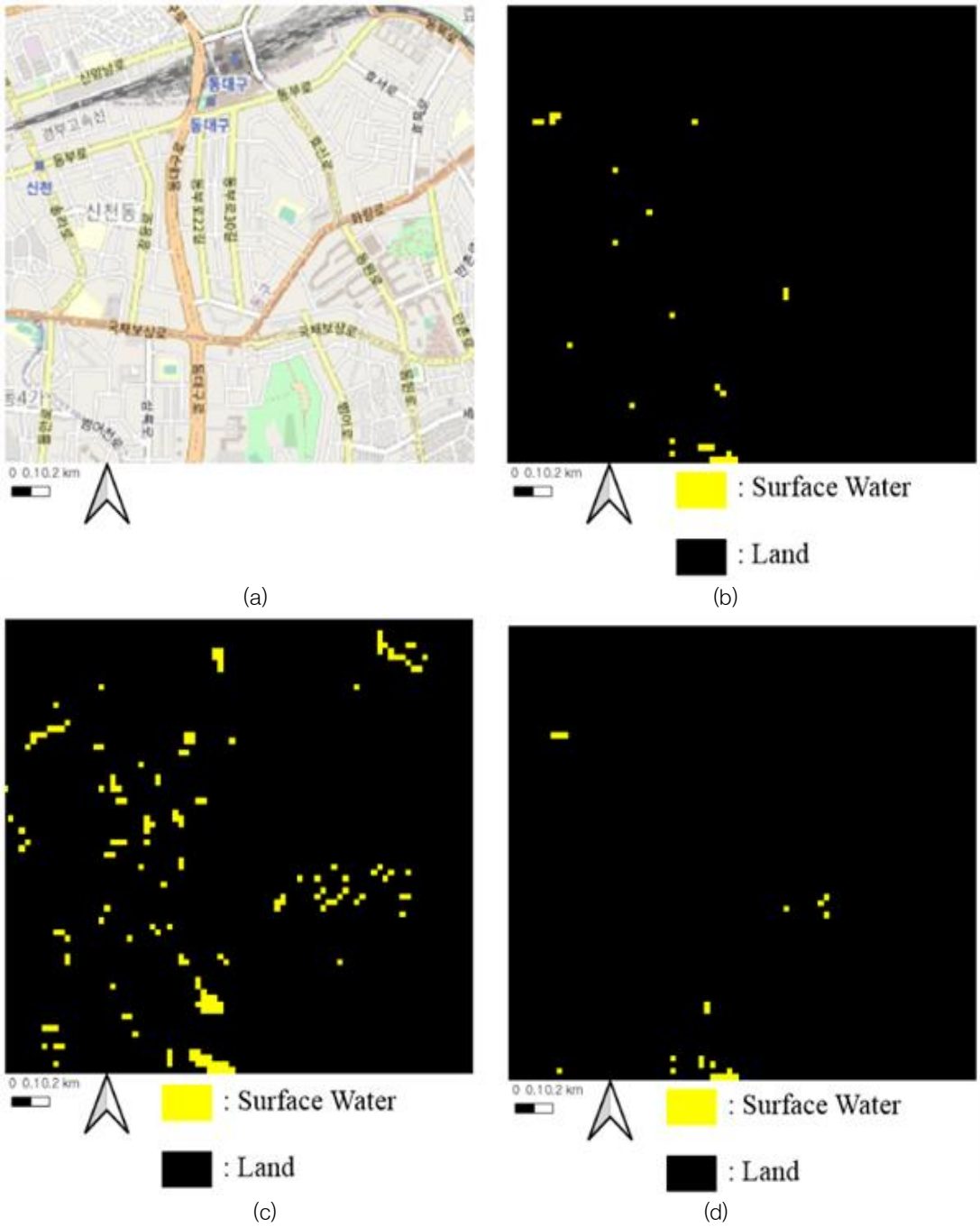


FIGURE 7. Examples of misclassification errors in the binary images: (a) Test bed shown in Open Street Map, (b) Binary image generated from NDWI, (c) Binary image generated from MNDWI, and (d) Binary image generated by SVM

능 기법인 SVM이 수분지수인 MNDWI에 비해 Landsat-8 위성영상으로부터 지표수를 탐지하는 데 상대적으로 더욱 우수하다는 사실을 알 수 있었다. 그러나 그림자 등의 원인으로 일부 육지 분류 결과들이 지표수 객체로 오분류 되었으며 이는 그림자의 영향을 받는 광학 위성영상의 한계점으로 파악되었다. 따라서 추후 연구에서는 그림자의 영향이 상대적으로 적은 LiDAR (Light Detection And Ranging) 점군 자료 등을 광학 위성영상과 함께 이용하거나 본 논문에서 제시한 세 가지 방법을 결합하는 등 다양한 접근법을 통해 지표수 탐지 정확도를 개선 시킬 필요가 있다. 본 연구에서는 Landsat-8 위성영상의 공간해상도(30m)로 인해 소규모 저수지 등 다양한 형태의 지표수를 탐지하는데 한계가 있으므로, 추후 연구에서는 Sentinel-2, KOMPSAT Series 등 고해상도 위성영상을 활용하여 하천의 분류 및 지류, 습지, 대규모 호수 및 저수지 등 다양한 형태의 지표수를 탐지하고 정확도를 검증할 계획이다. KAGIS

REFERENCES

Acharya, T.D., A. Subedi and D.H. Lee. 2018. Evaluation of Water Indices for Surface Water Extraction in a Landsat 8 Scene of Nepal. Sensors 18(8):2580.

Acharya, T.D., A. Subedi and D.H. Lee. 2019. Evaluation of Machine Learning Algorithms for Surface Water Extraction in a Landsat 8 Scene of Nepal. Sensors 19(12):2769.

Choung, Y.J. and M.H. Jo. 2016. Shoreline Change Assessment for Various Types of Coasts Using Multi-Temporal Landsat Imagery of the East Coast of South Korea. Remote Sensing Letters 7(1):91-100.

Choung, Y.J. and M.H. Jo. 2017. Comparison between a Machine-learning-based Method and a Water-index-based Method for Shoreline Mapping Using a High-Resolution Satellite Image Acquired in Hwado Island, South Korea. Journal of Sensors 2017:1-13.

Choung, Y.J. and J.M. Kim. 2019. Study of the Relationship between Urban Expansion and PM_{10} Concentration Using Multi-Temporal Spatial Datasets and the Machine Learning Technique: Case Study for Daegu, South Korea. Applied Sciences 9(6):1098.

Hu, H. 2006. Modification of Normalized Difference Water Index(NDWI) to Enhance Open Water Features in Remotely Sensed Imagery. International Journal of Remote Sensing 27(14):3025-3033.

Hwang, E.H., H.S. Chae and W.S. Yu. 2018. Development Plan of Compact Satellite for Water Resources and Water-related Disaster Management. Journal of the Korean Association of Geographic Information Studies 21(4):218-237 (황의호, 채효석, 유완식. 2018. 수자원·수재해 중형위성 개발 방안. 한국지리정보학회지 21(4):218-237).

Jo, M.H. 2012. A Study on the Extractiono of a River from the RapidEye Image Using ISODATA Algorithm. Journal of the Korean Association of Geographic Information Studies 15(4):1-14 (조명희. 2014. ISODATA 기법을 이용한 RapidEye 영상으로부터 하천의 추출에 관한 연구. 한국지리정보학회지 15(4):1-14).

Kim, D.Y., S.H. Baeck, G.H. Park, E.H. Hwang and H.S. Chae. 2017. Applicability of Water Resource Specialized Satellites for Observing Disasters on the Korean

Peninsula. Journal of the Korean Association of Geographic Information Studies 20(1): 85-97 (김동영, 백승협, 박광하, 황의호, 채효석. 2017. 한반도 수재해 관측을 위한 수자원 위성의 적용성. 한국지리정보학회지 20(1):85-97).

Kim, S. and Y. Lee. 2018. Present Status and Future Prospect of Satellite Image Uses in Water Resources Area. Korean Journal of Ecology and Environment 51(1):105-123 (김성준, 이용관. 2018. 수자원분야의 위성영상 활용 현황과 전망. 생태와 환경 51(1):105-123).

McFeeters, S.K. 1996. The Use of the Normalized Difference Water Index(NDWI) in the delineation of open water features. International Journal of Remote Sensing 17(7):1425-1432.

Microsoft. 2020. What is Machine Learning? https://azure.microsoft.com/ko-kr/overview/what-is-machine-learning-platform/. (Accessed December 2, 2020).

Moon, G.S., K.S. Kim and Y.J. Choung. 2020. Land Cover Classification Based on High Resolution KOMPSAT-3 Satellite Imagery Using Deep Neural Network Model. Journal of the Korean Association of Geographic Information Studies 23(3): 252-262 (문갑수, 김경섭, 정윤재. 2020. 심층신경망 모델을 이용한 고해상도 KOMPSAT-3 위성영상 기반 토지피복분류. 한국지리정보학회지 23(3):252-262).

OpenStreetMap. 2020. https://www.openstreetmap.org/#map=15/35.8207/128.6087. (Accessed December 2, 2020).

ScienceAll. 2020. Surface Water. https:// www.scienceall.com/%EC%A7%80%ED%91%9C%EC%88%98surface-water/. (Accessed December 1, 2020).

Shen, L. and C. Li. 2010. Water Body Extraction from Landsat ETM+ Imagery using Adaboost Algorithm. Proceedings of 18th International Conference on Geoinformatics, Beijing, China, pp.1-4.

USGS(United States Geological Survey). 2020. EarthExplorer. https://earthexplorer.usgs.gov/. (Accessed December 2, 2020).

Zhai, K., X. Wu, Y. Qin and P. Du. 2015. Comparison of Surface Water Extraction Performances of Different Classic Water Indices Using OLI and TM imageries in Different Situations. Geo-spatial Information Sciences 18(1):32-42. KAGIS